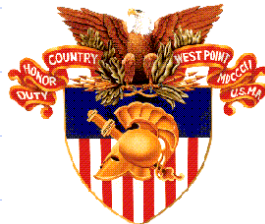


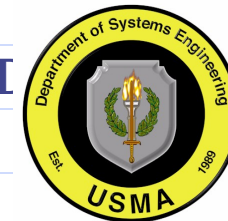


Validating Computational Human Behavior Models: Consistency and Accuracy Issues

**The MOVES Institute Open House
24-26 August 2004
Monterey, CA**

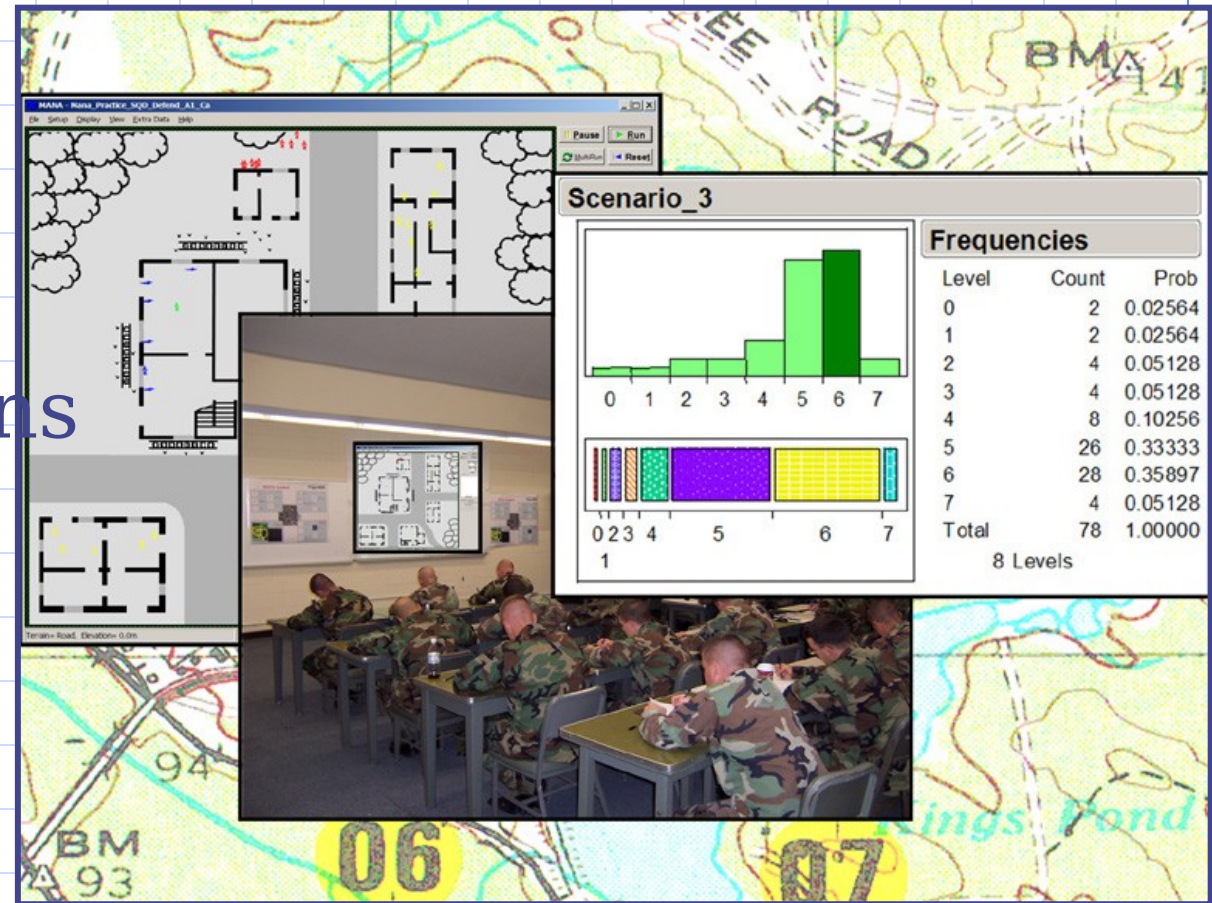


MAJ(P) Simon R. Goerger, Ph.D.
Simon.Goerger@usma.edu
(854) 938 - 5535



Outline

- Introduction
- Methodology
- Experiment
- Results
- Recommendations
- Contributions
- Conclusions
- Summary



Introduction:

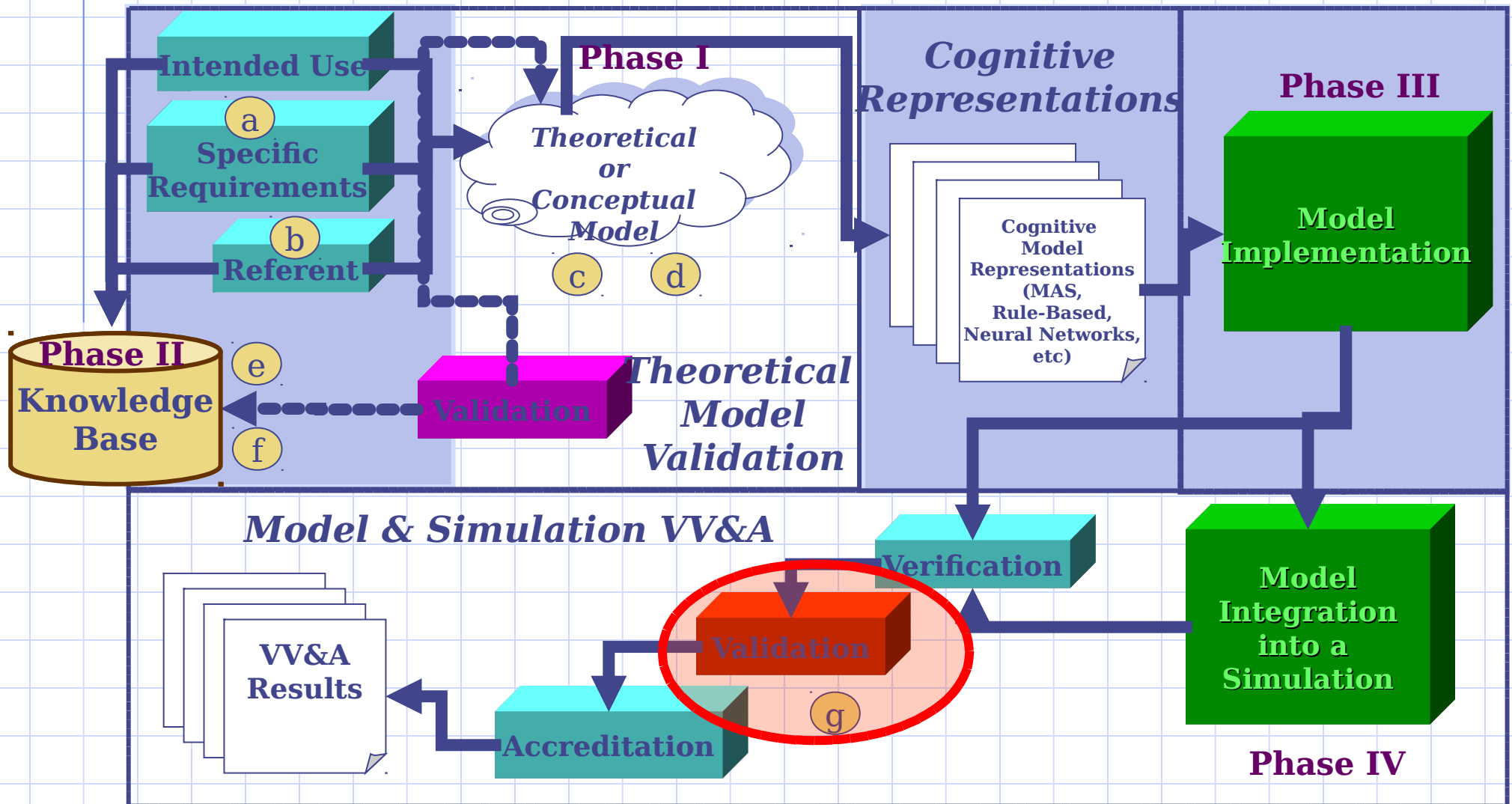
Motivation

- The world threat is changing and we can't focus solely on Soviet Tactics
- Viable human behavioral representation (HBR) models are needed for combat simulations
 - "Garbage in equals garbage out"
- Validation is hard for physically based models
 - The "how to" is not well defined
- Validation of HBR models is even more difficult
 - The nondeterministic nature of human cognitive processes
 - Inadequate quantitative measures for validating human behavior representation models
 - Subject Matter Expert (SME) bias

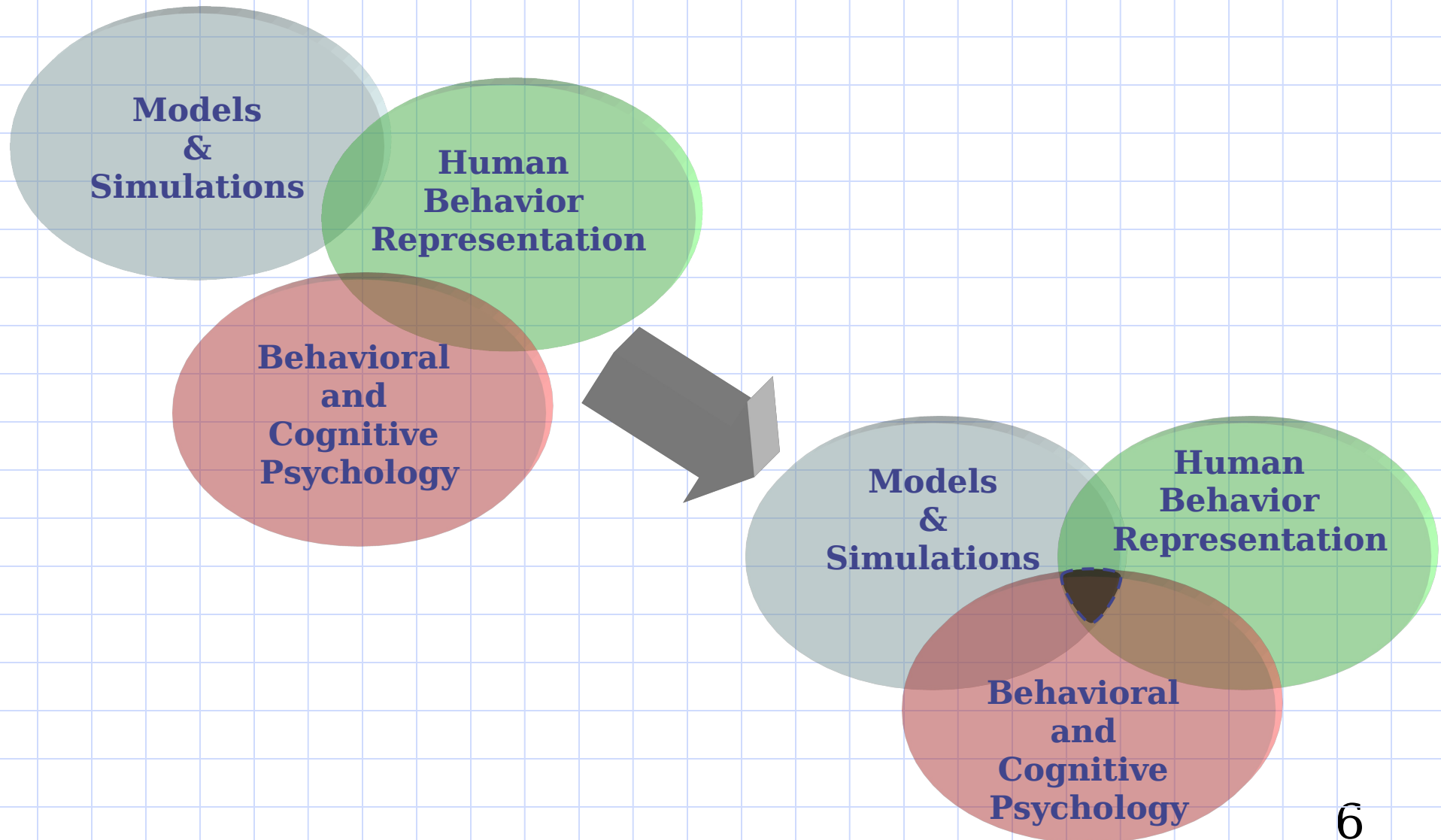
Introduction: Goal

To present a methodology for validating HBR model implementations for use in Department of Defense training and research models and simulations.

Introduction: Validation



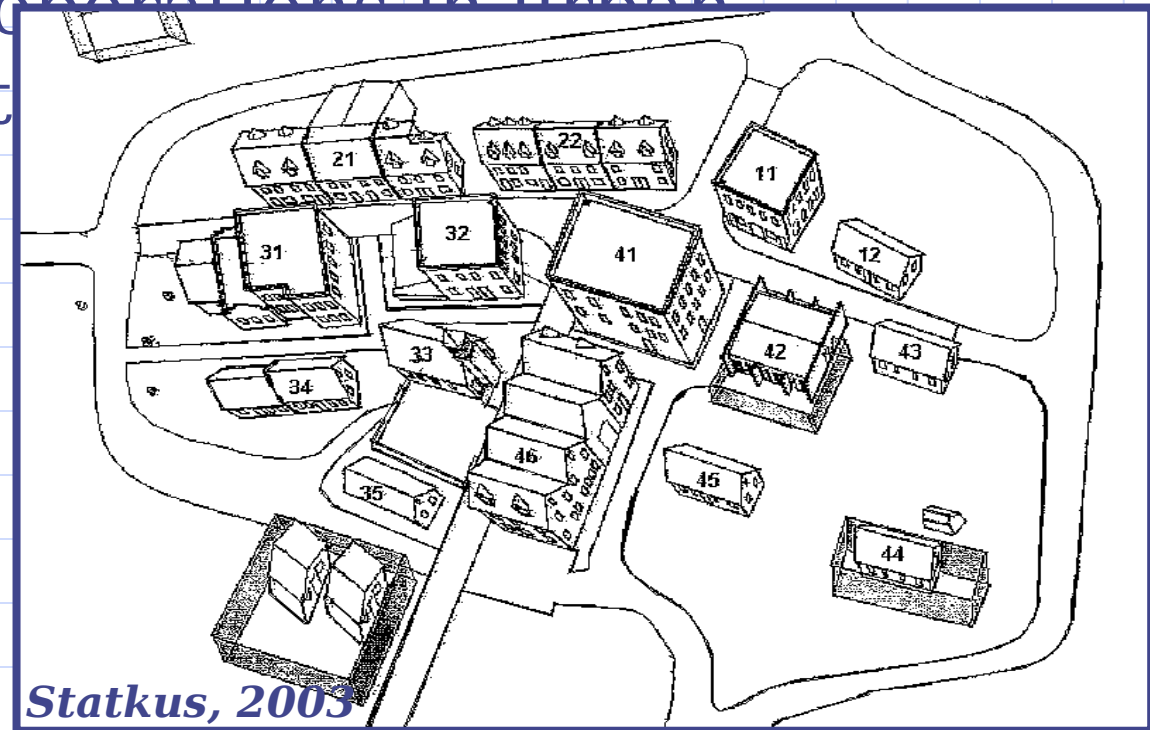
Methodology



Experiment: Simulation

Environment

- Map Aware Non-uniform Automata (MANA)
 - agent-based model
 - consists of entities representing military units
- McKenna military operations in urban terrain (MOUT) Site



Results: SME Bias

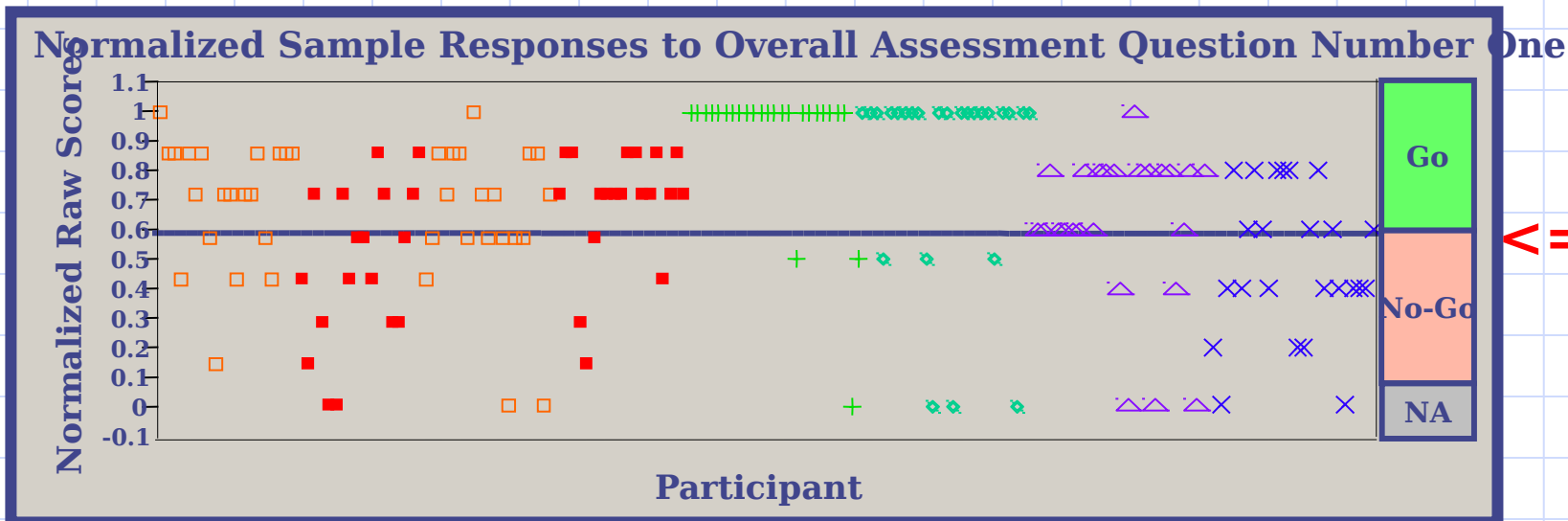
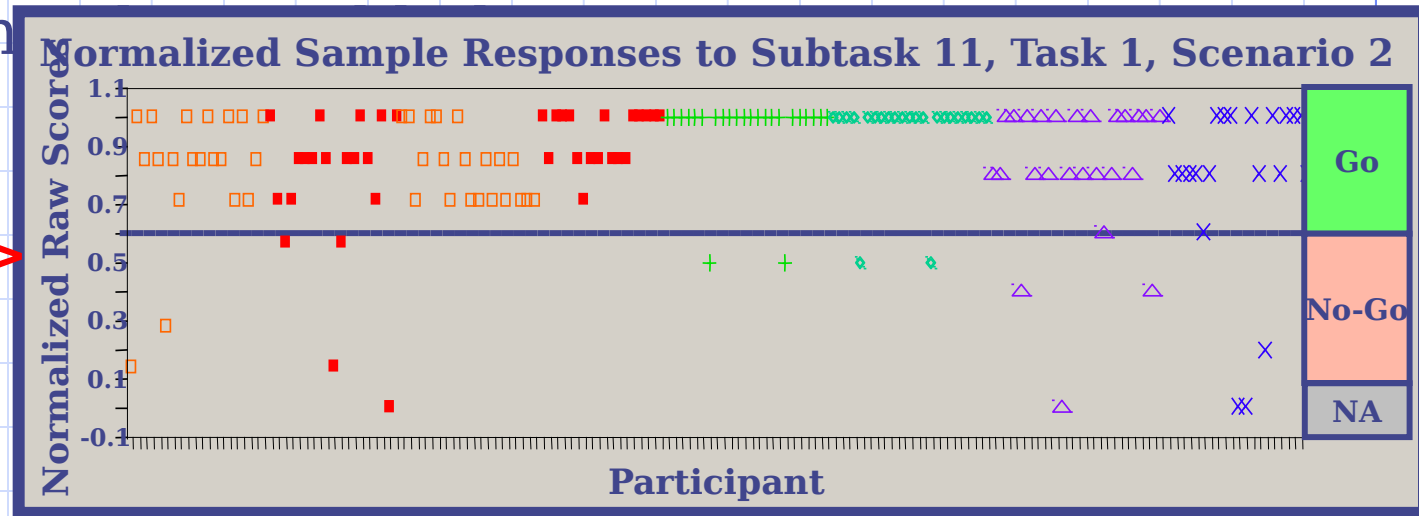
- “Systematic error introduced into sampling or testing by selecting or encouraging one outcome or answer over others” - Merriam-Webster’s Collegiate Dictionary
- Types
 - Performance
 - Anchoring
 - Contrast
 - Confirmation

Results: Inter-SME

Consistency

- 67% or more of the participants agree in their assessment of the
- Examples

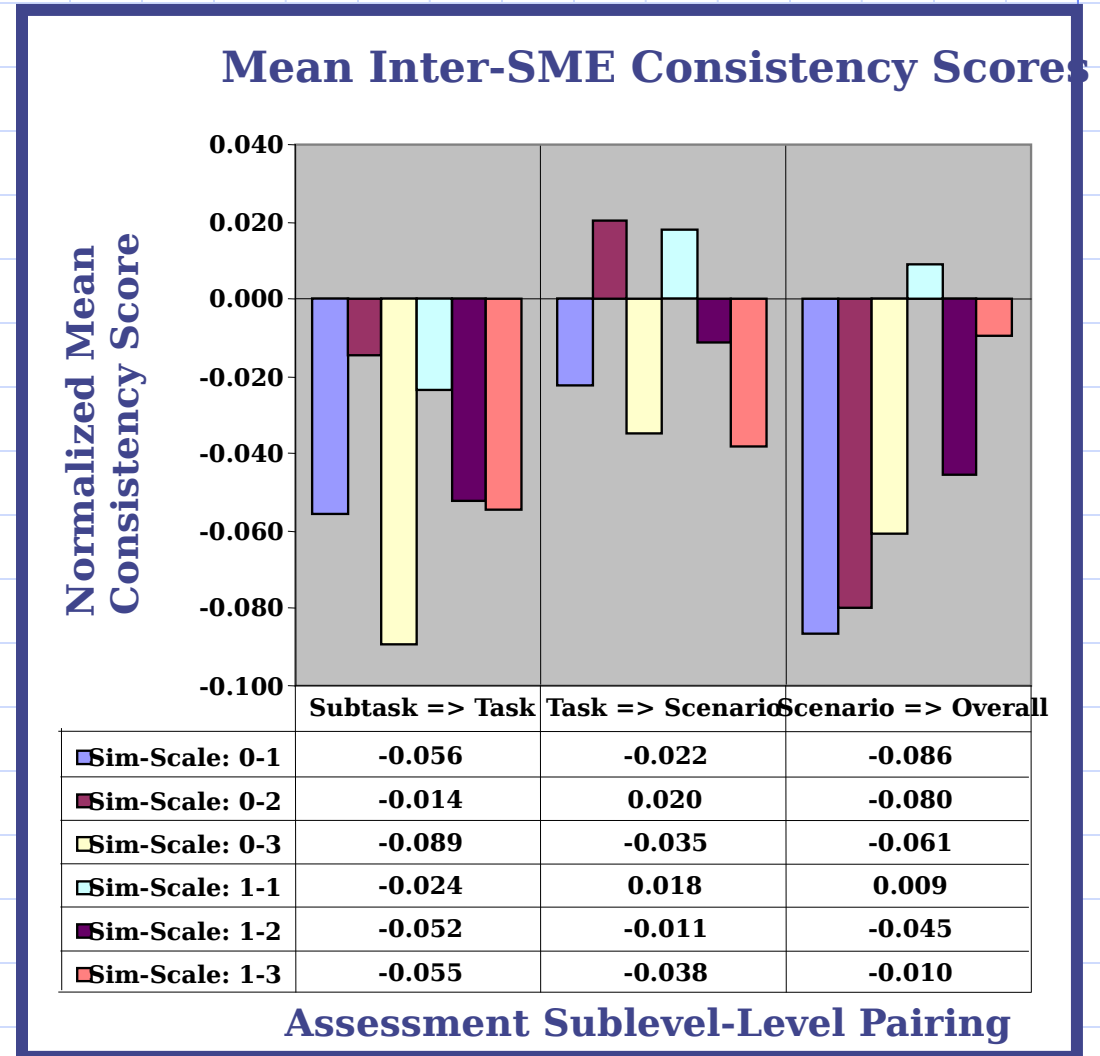
Consistent =>



<= Inconsistent

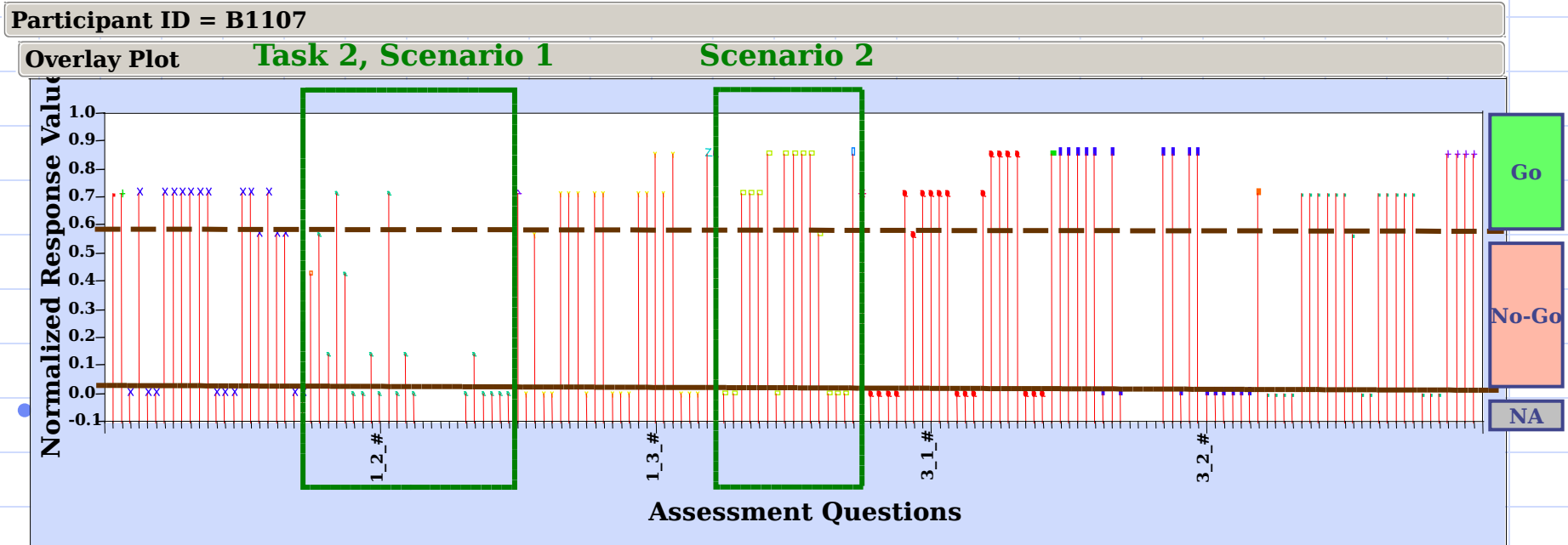
Results: Intra-SME

- **Consistency**
Participants were internally consistent in their assessments when their level score is the same as the mean score of their sublevel scores.
- **Results:** 150 (82.4 %) SMEs were internally inconsistent in their assessment.



Results: Anchoring Bias

- Participant assesses 90% or more of the questions after the primary task the same as they rated the initial task
- Pattern



Participant B1107 - Base Responses (Positive)

--- Undecided Line
— NA Response

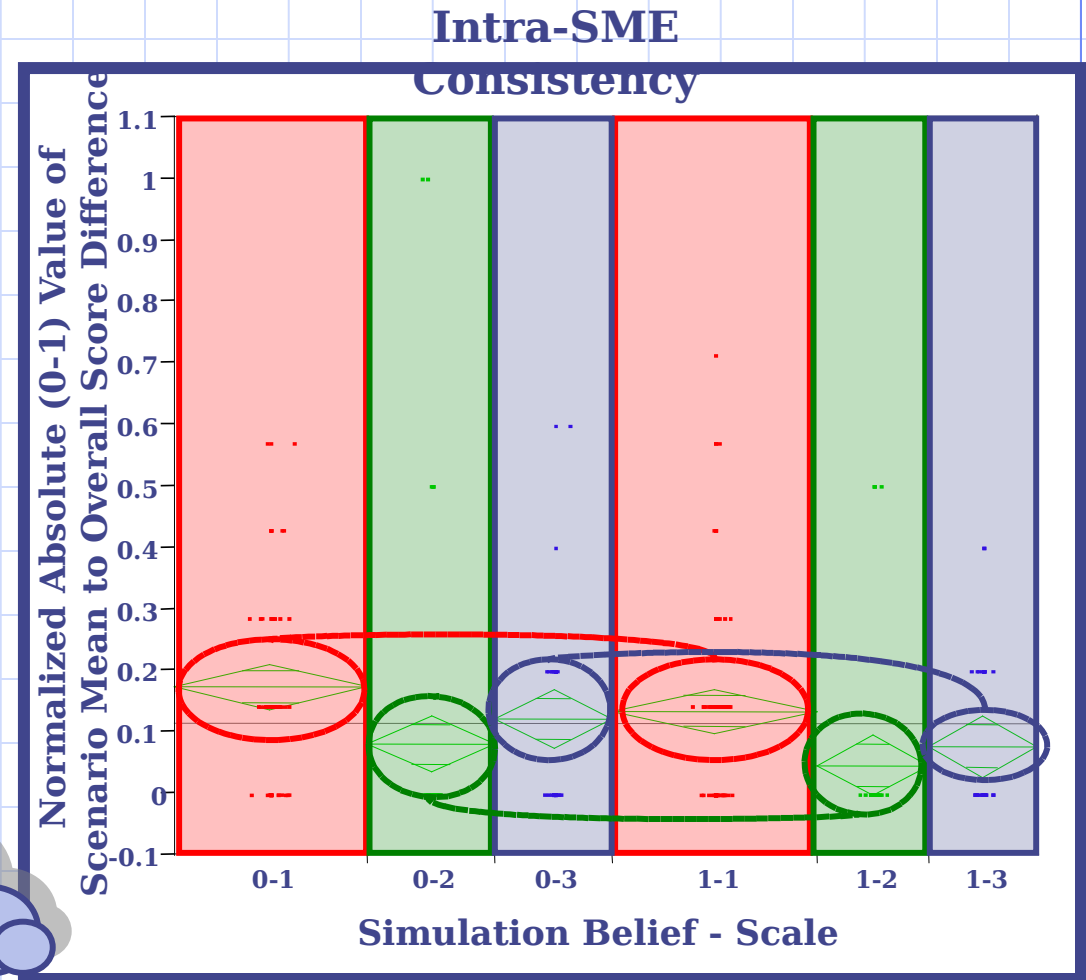
Results: Primary Hypothesis

H_o : SMEs demonstrate the same levels of effect on consistency and accuracy during validation of an HBR model implementation using a 7-Point Likert Scale as they do when using a 5-Point Likert Scale or Go/No-Go Scale.

Results: Primary Hypothesis

- Reject the null hypothesis and conclude there is effect on intra-SME consistency

The Go/No-Go Scale is the most consistent and the 7-Point Likert Scale is the least consistent.



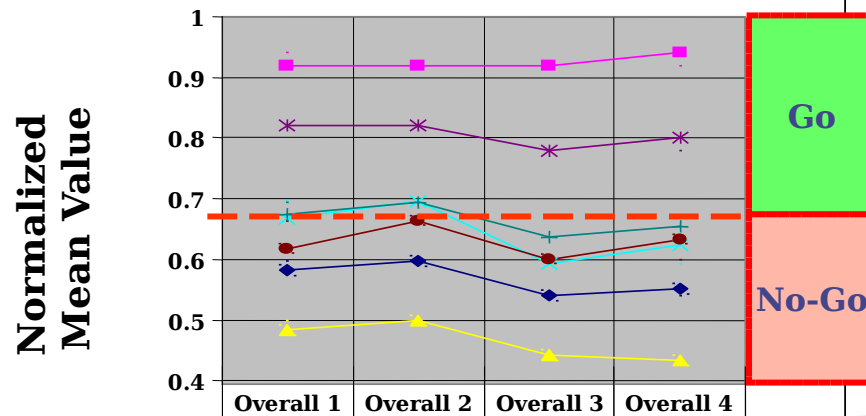
Simulation Belief: 0 - Live; 1 - Constructive

Scale: 1 - 7-Pt; 2 - Go/No-Go; 3 - 5-Pt

Results: SME Bias

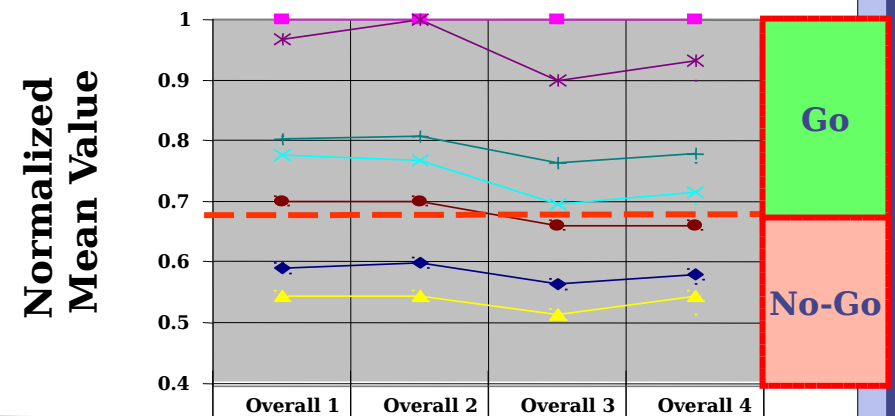
Inter-SME Consistency
Scores increase in all
cases but one, when bias is
removed.

SME Normalized Responses



Overall Assessment Question

SME Normalized Responses w/o Bias



Overall Assessment Question

Recommendations

- Training of SMEs
 - Ensure compliance with assessment process
 - Identify biases
- Assessment Scales
 - Grounding assessment scales with specific examples of performance
 - Weighting factors for assessment questions
- Automation
 - Identify bias patterns
 - Track and tally assessment scores
 - Prompt for justification of inconsistent responses

Contributions

- Lessons learned from the use of human behavior evaluation techniques, such as naturalistic decision-making, in the assessment of human behavior models
- Identified means to increase the consistency and accuracy of 'face validation' procedures for HBR models
- Formulation of new techniques for identifying and measuring the presence and impact of participant consistency and accuracy
- Identified quantitative patterns of bias based on SME responses to assessment questions
- Identified methods for removal of participant bias to mitigate participant inconsistencies and inaccuracies
- Establish a statistically significant relationship between bias and Neuroticism, Extraversion, and Openness Five-Factor Inventory personality styles
- Proposed a research agenda for the future enhancement of human behavior representation model validation procedures

Conclusions

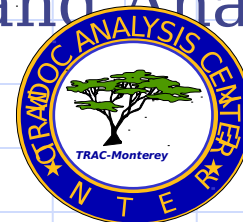
- Validating HBR models is difficult but must be done
- Human Performance Evaluation techniques provide statistically similar results when comparing human and simulated behavior; no difference in bias given simulation belief
- Assessment scales used shows effects on consistency and accuracy
- Subject matter expert biases show effects on consistency and accuracy

Summary

The DoD needs viable human behavior representation models for training and analysis. Validation of human behavior representation models is a difficult process which is not well defined. Issues such as subject matter expert bias affect consistency and accuracy of results. Assessment scales can mitigate inconsistency and inaccuracy.

Acknowledgments and Sponsors

- Navy Modeling & Simulation Management Office (NAVMSMO)
- Defense Modeling and Simulation Office (DMSO)
- Marine Corps Combat Development Command (MCCDC)
- US Army Natick Soldier Center
- Infantry Captains Career Course (ICCC)
- Training and Doctrine Command Analysis Center – Monterey (TRAC-Monterey)



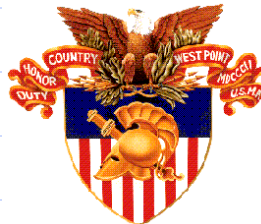


Questions?



Validating Computational Human Behavior Models: Consistency and Accuracy Issues

**The MOVES Institute Open House
24-26 August 2004
Monterey, CA**



MAJ(P) Simon R. Goerger, Ph.D.
Simon.Goerger@usma.edu
(854) 938 - 5535

